DOCUMENT RESUME

ED 292 818 TM 011 194

AUTHOR Roth, Rodney
TITLE The Differences between Teachers and Teacher Educators When Judging the NTE Professional Knowledge Test To Determine a Cut-Score.
PUB DATE Nov 87
NOTE 10p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (Mobile, AL, November 11-13, 1987).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Cutting Scores; Elementary Secondary Education; Higher Education; Item Analysis; Standardized Tests; Teacher Attitudes; *Teacher Certification; Teacher Education Curriculum; Teacher Educators; *Test Validity
IDENTIFIERS *National Teacher Examinations

ABSTRACT

A study involving 29 teachers and 21 teacher educators was conducted to determine whether differences occur between the two groups on a cut-score study of the National Teacher Examination (NTE) Professional Knowledge Test. The 50 professionals were the judges for a validation and cut-score study prior to implementing use of the NTE for initial certification in Arkansas. The teachers assessed each test item for its job-related relevance to determination of minimum teacher qualification and estimated the percent of minimally qualified candidates who would correctly answer the item. The teacher educators judged each item with respect to content coverage in the teacher education curriculum and made the same estimates as did teachers concerning the percent of minimally qualified candidates who would correctly answer the item. Results indicate that: (1) 91 of the 105 items were considered valid; (2) mean percent estimates for teachers were larger than those for teacher educators on 87 of the 91 items; and (3) the difference between the cut-score for teacher educators and the combined groups was quite small. (TJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED292818

THE DIFFERENCES BETWEEN TEACHERS AND TEACHER EDUCATORS WHEN JUDGING THE NTE PROFESSIONAL KNOWLEDGE TEST TO DETERMINE A CUT-SCORE

by

RODNEY ROTH, DEAN

COLLEGE OF EDUCATION

UNIVERSITY OF ALABAMA

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Rodney Roth

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

TM 011 194

Paper Presented at the 1987 Annual Meeting
of the
Mid-South Educational Research Association
November, 1987
Mobile, AL

BEST COPY AVAILABLE

One of the many decisions that has to be made for a NTE validation/cut-score study concerns what group or groups of judges to use in the study. ETS (1976) in their validity study for South Carolina used college faculty, typically teacher educators, as judges for the validation/cut-score study. This study compared the NTE test with the teacher education curriculum. In addition, the college faculty were used to determine the cut-score. This was the primary validation/cut-score model used from 1976-1981 in at least six states.

Roth (1982) in his NTE validation/cut-score study for Arkansas developed a totally new model. His study, based on the Uniform Guidelines on Employee Selection Procedures published by the Equal Employment Opportunity Commission in 1978, assessed the job relatedness or relevance of the NTE and not its relationship to the teacher education curriculum. He used teachers and teacher educators for the validity and cut-score judgments.

ETS (1983) developed a new model in 1983. They kept the original South Carolina procedures, but added the procedure developed by Roth. This current model uses teacher educators for the curriculum review. Teachers and teacher educators are used for the job relevance review and standard or cut-score judgments.

Poggio (1986) and others at the University of Kansas expanded, and in some ways modified, the "current" ETS model in their NTE study for Kansas. For example, they conducted an

equivalent forms validity study. In addition, they did not use teacher educators for the job relatedness part of the study and they used only teacher educators for the cut-score judgments.

PURPOSE OF STUDY

The purpose of this study is to determine if differences occur between teachers and teacher educators on a cut-score study of the NTE Professional Knowledge test. The two major questions addressed by the study are:

1. What are the differences between teachers and teacher educators on their item judgments of the percent of minimally qualified certification candidates who would correctly answer the item?; and
2. What cut-scores would be derived from teacher educator judges, teacher judges and a combination of teachers and teacher educators?

DATA COLLECTION

The data for this study were collected from 29 teachers and 21 teacher educators in December, 1986. These 50 professionals were the judges for a validation and cut-score study in order for the State of Arkansas to use the NTE Professional Knowledge test for initial certification.

The teachers were selected for the study based on two criteria. One criterion for the teacher selection process was to have teachers from school districts of varying size. The other criterion was certification field. Since the Professional Knowledge test was going to be used for all

teacher certification fields, teachers would need to represent the various teaching fields. Table 1 presents the certification fields for the 29 teachers.

TABLE 1

Public School Judges
Professional Knowledge

Certification Field	Number
English	1
Business Education	2
Social Studies	1
Science	1
Physical Education	5
Industrial Arts	1
Special Education	6
Music	1
Elementary Education	11

All of the colleges in Arkansas were asked to provide judges. The number of judges per college was determined by the proportion of teacher graduates from the various colleges from 1982-85.

The data collection session began with a short training session. It included the need for a state validation of the NTE in Arkansas and the NTE study design. It also included the written instructions that were given to each judge.

The teachers judged each test item for its job-related relevance to be a minimally qualified teacher and estimated

the percent of minimally qualified candidates who would correctly answer the item. The teacher educators judged each item with respect to content coverage in the teacher education curriculum and made the same estimates as teachers concerning the percent of minimally qualified candidates who would correctly answer the item.

DATA ANALYSIS PROCEDURES

The validity (relevance) of each item was determined by computing an item mean for each item on the relevance scale. This scale had a range from one (Not Relevant) to four (Crucial). In order for an item to be considered valid, the mean score on the relevance scale had to be greater than 2.5. In other words, the item had to be rated by the judges as closer to the important category than to the questionable category. If half of the judges had rated the item questionable and the other half had rated the item important, then the item would not have met the validity criterion since the mean rating would have been 2.50.

The validity (content coverage) of each item was determined by whether a majority of judges had responded that at least 90% of the students in their teacher training programs, regardless of certification field, would have had an opportunity to acquire the knowledge to answer the item correctly or at least 90% of the students would have had an opportunity to acquire the knowledge to answer the item correctly in high school or other experiences before their undergraduate program.

RESULTS

Ninety-one of the 105 items were considered valid. These 91 items met the relevance criterion from the judgments by teachers and the content coverage criterion from the judgments by teacher educators. A t-test comparing the mean percent estimated by teachers with the mean percent estimated by teacher educators was computed for each of the 91 items.

The results indicated that the mean percents for teachers were larger than for teacher educators on 87 of the 91 items. The four items larger for teacher educators were not significantly different. Nine items were significant at the .01 level and 14 items were significant at the .05 level. The remaining 68 items were not significantly different between the two groups.

A raw score cut-score is determined by adding the item mean percents for the valid items. The cut-score for teacher responses was 51 and for teacher educators was 44. The cut-score for the combined group was 48. The conversion formula from the raw score (R.S.) to the NTE scaled score is .8654 X R.S. + 600. The results for the three cut-scores are in Table 2.

TABLE 2

Professional Knowledge Cut-Scores

Group	Raw Score	Scaled Score	Percentile
Teachers	51	645	16
Teacher Educators	44	639	9
Combined	48	642	12

CONCLUSION

Since 1982, most NTE validation studies for initial state teacher certification have used teachers and teacher educators as judges to provide the data for the cut-score. Recently, Poggio (1986) and his colleagues at the University of Kansas, challenged the use of teachers in the standard setting process. They state:

> The standard setting method frequently employed (e.g., Angoff, Ebel) requires judgments about the expected performance on a test item of a referent group such as minimally competent beginning teachers. It seemed to us that college faculty were in a better position to make these judgments than were field based professionals. College faculty are more attuned to curricular and instructional considerations and to the knowledge and skill level of students graduating from teacher training programs. This insight qualifies them to make the judgments required by the standard setting methods. (Poggio, 1986)

The results of this study seem to indicate that the difference between the cut-score for teacher educators and the combined groups is quite small.

The real question, however, concerns defending the various decisions that are made during a validation and cut-score study in a court of law. In this case, the decision to use teachers and teacher educators instead of only teacher educators needs a reasonable rationale. Poggio (1986) provides a reasonable rationale for using teacher educators.

Courts tend to use what is called the "reasonable man standard" when they accept or reject testimony. In other words, courts would probably accept using teachers to set the standard if it appeared to be reasonable based on the rationale given during testimony.

One reasonable argument would be that teachers should be involved in deciding who should be able to enter the profession. Based on the results of this study, teachers have higher standards for admission to the profession than teacher educators.

One could also use the legal profession as an example. In most states, the decision for passing or failing the bar exam is determined by practicing lawyers and not law professors.

One additional reason (I am sure readers could add appropriate reasons) is that teachers do have an appropriate knowledge base to make the judgments. This knowledge base comes from their background, experience as teachers and their work with other teachers.

REFERENCES

Educational Testing Service. Report on a Study of the NTE Core Battery Tests by the State of New York. Princeton, NJ.: Educational Testing Service, September 1983.

Educational Testing Service. Report on a study of the use of the National Teacher Examinations by the state of South Carolina. Princeton, NJ.: Educational Testing Service, 1976.

Poggio, John P., Glasnapp, D., Miller, M., Tollefson, N., and Burry, J. "Strategies for Validating Teacher Certification Tests," Educational Measurement: Issues and Practice, Volume 5, Number 2, Summer 1986.

Roth, R. "Validation of the NTE: Arkansas Style." Paper presented at the Educational Commission of the States Annual Conference on Large-Scale Assesment, Boulder, Colorado, June 1982. Published in Resources in Education, March 1983.